AD-A102 562   MISSOURI UNIV-COLUMBIA DEPT OF STATISTICS   F/G 12/1
ORDER RESTRICTED STATISTICAL TESTS ON MULTINOMIAL AND POISSON P--ETC(U)
JUN 81   R L DYKSTRA, T ROBERTSON   N00014-78-C-0655
UNCLASSIFIED   TR-105   NL

1 of 1
AD A
102562

END
DATE
FILMED
9-81
DTIC

AD A102562

LEVEL II

University of Missouri-Columbia

Order Restricted Statistical Tests on Multinomial and Poisson Parameters: The Starshaped Restriction

by

Richard L. Dykstra
Department of Statistics, University of Missouri-Columbia

Tim Robertson
Department of Statistics, University of Iowa

Technical Report No. 105
Department of Statistics June 1981

Mathematical Sciences

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 105	2. GOVT ACCESSION NO. AD-A102562	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Order Restricted Statistical Tests on Multinomial and Poisson Parameters: The Starshaped Restriction.		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard L. Dykstra, Tim Robertson		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0655 ~~N00014-80-C-0321~~
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Stat., University of Missouri, Columbia, MO 65211; Dept. of Stat., University of Iowa, Iowa City, IA 52240		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Department of the Navy, Arlington, VA		12. REPORT DATE June 1981
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

(6) (9) (10) (15) (11) (12) 27

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution unlimited

14/ TR-105

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE AUG 7 1981 S C

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Starshaped ordering, Multinomial, Poisson, Likelihood ratio tests, chi-bar-square distributions, maximum likelihood.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Likelihood ratio statistics for (i) testing the homogeneity of a collection of multinomial parameters against the alternative which accounts for the restriction that those parameters are starshaped (cf. Shaked, Ann. Statist. (1979)), and for (ii) testing the null hypothesis that this parameter vector is starshaped are considered. For both tests the asymptotic distribution of the test statistic under the null hypothesis is a version of the chi-bar-square distribution. Analogous tests on a collection of Poisson means are also found to have asymptotic chi-bar-square distributions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

400660

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ORDER RESTRICTED STATISTICAL TESTS ON MULTINOMIAL AND POISSON PARAMETERS: THE STARSHAPED RESTRICTION

Richard L. Dykstra* and Tim Robertson

(University of Missouri, Columbia and University of Iowa)

This work was supported by ONR Contracts N00014-78-0655 and N00014-80-C-0321.

*This work was done while on sabbatical leave at The University of Iowa, 1980-1981.

ABSTRACT

Likelihood ratio statistics for (i) testing the homogeneity of a collection of multinomial parameters against the alternative which accounts for the restriction that those parameters are starshaped (cf. Shaked, Ann. Statist. (1979)), and for (ii) testing the null hypothesis that this parameter vector is starshaped, are considered. For both tests the asymptotic distribution of the test statistic under the null hypothesis is a version of the chi-bar-square distribution. Analogous tests on a collection of Poisson means are also found to have asymptotic chi-bar-square distributions.

AMS 1970 Subject Classifications: Primary 62F05

Secondary 62E15

Keywords and phrases: starshaped orderings, multinomial Poisson, likelihood ratio tests, chi-bar-square distributions, maximum likelihood

1. Introduction and Summary.

Shaked (1979) derived the maximum likelihood estimate of a vector of Poisson (normal) means subject to the restriction that this vector is "starshaped." A vector $\underset{\sim}{\theta} = (\theta_1,\theta_2,\cdots,\theta_k)$ is said to be lower starshaped provided

$\theta_1 \geq \frac{\theta_1+\theta_2}{2} \geq \cdots \geq \frac{\theta_1+\theta_2+\cdots+\theta_k}{k} \geq 0$ with an analogous restriction defining an upper starshaped vector. Starshaped vectors arise naturally in reliability theory as well as in certain situations where finite populations are amalgamated. We refer the interested reader to Shaked (1979) for examples of parameter sets which might be known or suspected to satisfy such a restriction.

In Section 2 we consider a sampling situation where the result of each trial of our experiment must be a member of a set of mutually exclusive events with corresponding probabilities $p_1,p_2,\cdots,p_k$. The maximum likelihood estimate of the vector, $\underset{\sim}{p} = (p_1,p_2,\cdots,p_k)$, subject to the restriction that it be lower starshaped, is derived (this sampling situation was not considered by Shaked (1979)). This derivation is quite direct and elegant in light of the complexity of the analysis in Shaked (1979) and in light of the difficulties involved in the related problem of finding the maximum likelihood estimate of $\underset{\sim}{p}$ subject to the restriction $p_1 \geq p_2 \geq \cdots \geq p_k$ (cf. Barlow, Bartholomew, Bremner and Brunk (1972)). In fact, proofs that various algorithms for the solution to the latter problem yield the desired result are usually by induction.

In addition, asymptotic distribution theory for the likelihood ratio test of the homogeneity of $p_1, p_2, \cdots, p_k$ against the alternative that $\underset{\sim}{p}$ is starshaped and for testing that $\underset{\sim}{p}$ is is starshaped as a null hypothesis is also presented in Section 2. Again the derivations are relatively direct. In both situations, the tail probabilities under the null hypothesis of this asymptotic distribution turn out to be of the form

$$\bar{\chi}^2_{k-1}(t) = \Sigma_{\ell=0}^{k-1}\binom{k-1}{\ell-1}\left(\frac{1}{2}\right)^{k-1} P[\chi^2_\ell \geq t]$$

where χ^2_ℓ denotes a standard chi-square variable with ℓ degrees of freedom. A somewhat similar distribution is encountered in the problem of testing homogeneity when the alternative is restricted by $p_1 \geq p_2 \geq \cdots \geq p_k$ (cf. Chacko (1966)) and for testing $p_1 \geq p_2 \geq \cdots \geq p_k$ as a null hypothesis (cf. Robertson (1978) and related results in Robertson and Wegman (1978)). Such weighted chi-square distributions are encountered in many order restricted inference problems (cf. Barlow et al. (1972)). They were first encountered by Bartholomew (1959) and are usually called chi-bar-square distributions $(\bar{\chi}^2)$.

In Section 3 we assume independent samples from each of k Poisson populations. The analysis in Section 2 together with the well known fact that the joint distribution of independent Poisson random variables conditioned on the value of the sum multinomial, is used to derive maximum likelihood estimates under the starshaped restriction on the parameter values. This derivation is substantially easier than the original derivation of these estimates by Shaked (1979). Asymptotic distribution theory for likelihood ratio statistics used for testing homogeneity versus starshaped and for testing starshaped as a null hypothesis is also presented.

2. Multinomial Problem.

Suppose we have n independent trials of an experiment, the outcome of which must be one of k mutually exclusive events with corresponding probabilities, $p_1, p_2, \cdots, p_k$ $(\Sigma_{i=1}^k p_i = 1)$. Our first task is to find the maximum likelihood estimate of the vector $\underset{\sim}{p}$ under the restriction:

$$(2.1) \qquad H_1 : p_1 \geq \frac{p_1+p_2}{2} \geq \cdots \geq \frac{p_1+p_2+\cdots+p_{k-1}}{k-1} \geq \frac{1}{k} .$$

We define a one to one transformation of the parameter space by introducing new parameters $\theta_1, \theta_2, \cdots, \theta_{k-1}$ where

$$(2.2) \qquad \theta_i = (\Sigma_{j=1}^{i} p_j)/(\Sigma_{j=1}^{i+1} p_j) : i = 1,2,\cdots,k-1$$

$(p_1 = \Pi_{j=1}^{k-1}\theta_j,\ p_i = (1-\theta_{i-1})\Pi_{j=i}^{k-1}\theta_j;\ i = 2,3,\cdots,k-1,\ p_k = (1-\theta_{k-1}))$.

In terms of the θ's the likelihood function can be written

$$(2.3)\quad L(\underset{\sim}{\theta}) = \prod_{i=1}^{k-1} \theta_i^{n\Sigma_{j=1}^{i}\hat{p}_j}(1-\theta_i)^{n\hat{p}_{i+1}}, \quad 0 \le \theta_i \le 1,$$

where $\hat{p}_i$ is the relative frequency of the event having probability $p_i : i = 1,2,\cdots,k$. The restriction (2.1) becomes

$$H_1' : \theta_i \ge i/(i+1) : i = 1,2,\cdots,k-1.$$

It is easy to find the maximum of the function $\theta^a(1-\theta)^b$ subject to $\theta \ge c$ $(0 \le \theta \le 1)$. This maximum is attained at $\bar{\theta} = a/(a+b) \vee c$, where $\vee$ denotes the larger of the two numbers. It follows that the maximum likelihood estimates which satisfy H_1 are given by

$$(2.4)\qquad \bar{\theta}_i = \hat{\theta}_i \vee (i/i+1) : i = 1,2,\cdots,k-1$$

where $\hat{\theta}_i = (\Sigma_{j=1}^{i}\hat{p}_j)/(\Sigma_{j=1}^{i+1}\hat{p}_j)$ and

$$(2.5)\qquad \bar{p}_i = (1-\bar{\theta}_{i-1}) \prod_{j=i}^{k-1} \bar{\theta}_j \ ; \ i = 1,2,\cdots,k-1$$

with $\bar{\theta}_0 = 0$ and $\bar{p}_k = 1-\bar{\theta}_{k-1}$. Restating (2.5), we have established the following theorem.

<u>Theorem 2.1</u>. The maximum likelihood estimates of $p_1,p_2,\cdots,p_k$ subject to the lower starshaped restriction are given by

$$\bar{p}_i = \left[1 - \left(\frac{\Sigma_{j=1}^{i-1}\hat{p}_j}{\Sigma_{j=1}^{i}\hat{p}_j} \vee \frac{i-1}{i}\right)\right] \cdot \Pi_{j=i}^{k-1}\left(\frac{\Sigma_{\ell=1}^{j}\hat{p}_\ell}{\Sigma_{\ell=1}^{j+1}\hat{p}_\ell} \vee \frac{j}{j+1}\right); \quad i = 1,2,\cdots,k-1$$

and

$$\bar{p}_k = 1 - \left(\Sigma_{j=1}^{k-1}\hat{p}_j \vee \frac{k-1}{k}\right)$$

where $\hat{p}_i$ denotes the relative frequency of the event having probability p_i.

Turning to the testing problem, let

$$\text{(2.6)} \qquad H_0 : p_1 = p_2 = \cdots = p_k = 1/k.$$

We let Λ_{01} denote the likelihood ratio test statistic for testing H_0 against $H_1 - H_0$ (i.e., H_1 but not H_0) and let $T_{01} = -2 \ln \Lambda_{01}$. It is convenient to write T_{01} in terms of the θ's as follows:

$$\text{(2.7)} \qquad T_{01} = 2\Sigma_{i=1}^{k-1}\left\{\left(n\Sigma_{j=1}^{i}\hat{p}_j\right)[\ln \bar{\theta}_i - \ln(i/i+1)] + n\hat{p}_{i+1}[\ln(1-\bar{\theta}_i) - \ln(1/i+1)]\right\}.$$

Using Taylor's Theorem with a second degree remainder term, we expand $\ln \bar{\theta}_i$ and $\ln(i/i+1)$ about $\hat{\theta}_i$, and expand $\ln(1-\bar{\theta}_i)$ and $\ln(1/i+1)$ about $1-\hat{\theta}_i$. The linear terms drop out and we obtain

$$(2.8)\qquad T_{01} = 2\sum_{i=1}^{k-1}\left[-\frac{n\sum_{j=1}^{i}\hat{p}_j}{2\alpha_i^2}(\bar{\theta}_i-\hat{\theta}_i)^2 + \frac{n\sum_{j=1}^{i}\hat{p}_j}{2\beta_i^2}(\hat{\theta}_i-\frac{i}{i+1})^2 - \frac{n\hat{p}_{i+1}}{2\nu_i^2}(\hat{\theta}_i-\bar{\theta}_i)^2 + \frac{n\hat{p}_{i+1}}{2\gamma_i^2}(\hat{\theta}_i-\frac{i}{i+1})^2\right]$$

where α_i is between $\bar{\theta}_i$ and $\hat{\theta}_i$; β_i is between $\hat{\theta}_i$ and $i/(i+1)$; ν_i is between $(1-\bar{\theta}_i)$ and $(1-\hat{\theta}_i)$ and γ_i is between $(1-\hat{\theta}_i)$ and $1/(i+1)$. The law of large numbers implies that, under H_0, $\hat{\theta}_i$ converges to $i/i+1$.

To obtain some insight into the asymptotic power of the likelihood ratio test, we consider a sequence of alternatives $\underset{\sim}{p}_n$ satisfying H_1 which converges to $\underset{\sim}{p}$ where $p_i > 0$ for all i. We let $\hat{\underset{\sim}{p}}_n$ denote a random vector corresponding to $\underset{\sim}{p}_n$, i.e., $n\hat{\underset{\sim}{p}}_n$ is multinomial $(n,\underset{\sim}{p}_n)$. In our alternative parameterization we have

$$\theta_{n,i} = \sum_{j=1}^{i} p_{n,j} / \sum_{j=1}^{i+1} p_{n,j} \longrightarrow \theta_i = \sum_{1}^{i} p_j / \sum_{1}^{i+1} p_j$$

and

$$\hat{\theta}_{n,i} = \sum_{j=1}^{i} \hat{p}_{n,j} / \sum_{j=1}^{i+1} \hat{p}_{n,j}.$$

Somewhat surprisingly, it can be shown by conditioning on $\sum_{j=1}^{i+1} \hat{p}_{nj}$, that $E(\hat{\theta}_{n,i}) = \theta_{n,i}$.

Since $n\hat{\underset{\sim}{p}}_n$ is multinomial $(n,\underset{\sim}{p}_n)$, it can be shown that

$\sqrt{n}(\hat{\underset{\sim}{p}}_n - \underset{\sim}{p}_n) \xrightarrow{D} \mathrm{MVN}(\underset{\sim}{0}, \Sigma)$ where $\Sigma = (\sigma_{ij})$ is defined by

$$\sigma_{ij} = \begin{cases} p_i(1 - p_i), & i = j \\ -p_i p_j, & i \neq j. \end{cases}$$

(One way of showing this is to verify that the moments of linear combinations of $\sqrt{n}(\hat{p}_{n,i} - p_{n,i})$ converge to the moments of linear combinations of an appropriate MVN vector, and then employ Theorem B of Serfling (1980) and the Cramer-Wold device.)

Moreover, if we define the function $\underset{\sim}{g} = (g_1, \cdots, g_{k-1}): R^k \longrightarrow R^{k-1}$ by

$$g_i(\underset{\sim}{x}) = \sum_1^i x_j / \sum_1^{i+1} x_j,$$

Theorem 4.2 of Kepner (1979) implies that

$$\sqrt{n}[\underset{\sim}{g}(\hat{\underset{\sim}{p}}_n) - \underset{\sim}{g}(\underset{\sim}{p}_n)] = \sqrt{n}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_n) \xrightarrow{D} \mathrm{MVN}(\underset{\sim}{0}, D\Sigma D')$$

where

$$D = \left(\frac{\partial g_i}{\partial x_j}\right)^{k-1 \times k}_{\underset{\sim}{x} = \underset{\sim}{p}}$$

$$= \begin{pmatrix} \dfrac{p_2}{\left(\sum_1^2 p_j\right)^2} & \dfrac{-p_1}{\left(\sum_1^2 p_j\right)^2} & 0 & \cdots\cdots & 0 \\ \dfrac{p_3}{\left(\sum_1^3 p_j\right)^2} & \dfrac{p_3}{\left(\sum_1^3 p_j\right)^2} & \dfrac{-\sum_1^2 p_j}{\left(\sum_1^3 p_j\right)^2} & 0 \quad \cdots & 0 \\ \vdots & \vdots & & & \\ \dfrac{p_k}{\left(\sum_1^k p_j\right)^2} & \dfrac{p_k}{\left(\sum_1^k p_j\right)^2} & \cdots\cdots & \cdots\cdots & \dfrac{-\sum_1^{k-1} p_j}{\left(\sum_1^k p_j\right)^2} \end{pmatrix}.$$

Careful calculation reveals that $D\Sigma D' = (\psi_{ij})$ is given by

$$\psi_{ij} = \begin{cases} \dfrac{p_{i+1}\left(\sum_1^i p_j\right)}{\left(\sum_1^{i+1} p_j\right)^3}, & i=j \\ 0, & i \neq j, \end{cases} \tag{2.9}$$

so that, fortunately, we have asymptotic independence among the $\sqrt{n}\ \hat{\theta}_{n,i}$'s.

If we recall that $\bar{\theta}_{n,i} = \hat{\theta}_{n,i} \vee \frac{i}{i+1}$, we can express the likelihood ratio test statistic as

$$T_{01}^{(n)} = \sum_{i=1}^{k-1} (X_{n,i}+\delta_{n,i})^2 a_{n,i} - (X_{n,i}+\delta_{n,i})^2 b_{n,i} I[X_{n,i}+\delta_{n,i} \leq 0] \tag{2.10}$$

where

$$X_{n,i} = \sqrt{n}(\hat{\theta}_{n,i}-\theta_{n,i})\left[\frac{(i+1)^3}{ik}\right]^{1/2},$$

$$\delta_{n,i} = \sqrt{n}\left(\theta_{n,i} - \frac{i}{i+1}\right)\left[\frac{(i+1)^3}{ik}\right]^{1/2},$$

(2.11)

$$a_{n,i} = \left[\sum_{j=1}^{i} \frac{\hat{p}_{n,j}}{\beta_{n,i}^2} + \frac{\hat{p}_{n,i+1}}{\gamma_{n,i}^2}\right]\left[\frac{(i+1)^3}{ik}\right]^{-1}, \quad \text{and}$$

$$b_{n,i} = \left[\sum_{j=1}^{i} \frac{\hat{p}_{n,j}}{\alpha_{n,i}^2} + \frac{\hat{p}_{n,i+1}}{\nu_{n,i}^2}\right]\left[\frac{(i+1)^3}{ik}\right]^{-1}.$$

Since $\mathcal{P}_n$ satisfies H_1 for all n, $\delta_{n,i} \geq 0$ for all n,i. If $\delta_{n,i} \longrightarrow \delta_i$ (finite) as $n \longrightarrow \infty$, then $\theta_{n,i} \longrightarrow \frac{i}{i+1}$, so that $\hat{\theta}_{n,i} \xrightarrow{p} \frac{i}{i+1}$. Recalling how $\alpha_{n,i}$, $\beta_{n,i}$, $\gamma_{n,i}$, and $\nu_{n,i}$ were obtained, it then follows that $a_{n,i} \xrightarrow{p} 1$ and $b_{n,i} \xrightarrow{p} 1$ as $n \longrightarrow \infty$.

In this situation, $X_{n,i} \xrightarrow{\mathcal{D}} Z_i$, where Z_i is a $n(0,1)$ random variable. Using Theorem 4.9 of Billingsley (1968), we have

$$(X_{n,i} + \delta_{n,i}, a_{n,i}, b_{n,i}) \xrightarrow{\mathcal{D}} (Z_i + \delta_i, 1, 1).$$

Then noting that the function $h : R^3 \longrightarrow R^1$ defined by

$$h(x,y,z) = x^2 y - x^2 z \; I_{[x \leq 0]}$$

is continuous, we may use Theorem 5.1 of Billingsley to say

$$(X_{n,i}+\delta_{n,i})^2 a_{n,i} - (X_{n,i}+\delta_{n,i})^2 b_{n,i} \; I_{[X_{n,i}+\delta_{n,i} \leq 0]}$$

$$\xrightarrow{\mathcal{D}} (Z_i+\delta_i)^2 - (Z_i+\delta_i)^2 \; I_{[Z_i+\delta_i \leq 0]}$$

$$= [(Z_i+\delta_i) \vee 0]^2.$$

In the event that $\delta_{n,i} \longrightarrow \infty$, it can be shown that $a_{n,i}$ is bounded away from zero asymptotically while $X_{n,i}$ converges in distribution. Thus

$$(X_{n,i}+\delta_{n,i})^2 a_{n,i} - (X_{n,i}+\delta_{n,i})^2 b_{n,i}\, I_{[X_{n,i}+\delta_{n,i}\le 0]} \xrightarrow{P} \infty.$$

We have thus established the following theorem.

<u>Theorem 2.2</u>. If $\underset{\sim}{p}_n$ satisfying H_1 converges to $\underset{\sim}{p}(p_i>0$ for all i), and if $\delta_{n,i}$ (as defined in (2.11)) $\longrightarrow \delta_i$ (possibly ∞) for $i=1,\cdots,k-1$, then $T_{0,1}$ is distributed asymptotically as

$$U = \sum_{i=1}^{k-1} [(Z_i+\delta_i) \vee 0]^2, \tag{2.12}$$

where $Z_1,\cdots,Z_{k-1}$ are independent $n(0,1)$ random variables.

Of course the distribution of the random quantity in (2.12) is very intractable, except under the null distribution H_0 ($\delta_i=0$, $i=1,\cdots,k-1$) when it becomes surprisingly nice. To elaborate, suppose I is a subset of $\{1,2,\cdots,k-1\}$ and let E_I be the event $E_I = [Z_i \ge 0;\ i \in I$ and $Z_i < 0;\ i \notin I]$. Then, for any real number u,

$$\begin{aligned} P[U \ge u, E_I] &= P[\textstyle\sum_{i\in I} Z_i^2 \ge u,\ Z_i \ge 0;\ i \in I,\ Z_i < 0;\ i \notin I] \\ &= P[\textstyle\sum_{i\in I} Z_i^2 \ge u,\ Z_i \ge 0;\ i \in I]\cdot P[Z_i < 0;\ i \notin I] \\ &= P[\textstyle\sum_{i\in I} Z_i^2 \ge u \mid Z_i \ge 0;\ i \in I]\cdot (1/2)^{k-1} \\ &= P[\chi_m^2 \ge u]\cdot (1/2)^{k-1} \end{aligned}$$

where m is the number of elements in I. The last step follows from Lemma B on page 128 of Barlow, Bartholomew, Bremner and Brunk (1972). Partitioning the event $[U \geq u]$ by intersecting it with all such events, E_I, we obtain the expression for $P[U \geq u]$ given in the following theorem.

<u>Theorem 2.3</u>. If H_0 is true then

$$\lim_{n\to\infty} P[T_{01} \geq t] = \Sigma_{\ell=0}^{k-1}\binom{k-1}{\ell}(1/2)^{k-1}\; P[\chi_\ell^2 \geq t] = \bar{\chi}_{k-1}^2(t)$$

for all real t $(\chi_0^2 \equiv 0)$.

The expression (2.12) is useful in getting a feeling for the asymptotic power of the test compared to the usual unrestricted test for homogeneity of multinomial parameters. In particular, under the conditions of Theorem 2.2, the asymptotic distribution of $-2 \ln \Lambda$ where Λ is the unrestricted likelihood ratio (or of the usual Pearson chi-square goodness of fit test) is the same as

$$U' = \sum_1^{k-1} (Z_i + \delta_i)^2$$

where $Z_1, \cdots, Z_{k-1}$ are independent $n(0,1)$ random variables. Under H_0, U' is $\chi^2(k-1)$ and hence by looking at Theorem 2.3, we see that its critical point must be substantially larger than for testing H_0 against H_1. However, as the δ_i's become larger, more nonzero terms enter into (2.12), so that U and U' become more nearly equivalent. The smaller critical point of the restricted test implies that its power must be larger eventually than that of the unrestricted test.

We now turn to the problem of testing H_1 as a null hypothesis. Since the unrestricted maximum likelihood estimate of θ_i is equal to $\hat{\theta}_i$, it follows directly by writing the likelihood ratio in terms of $\hat{\underset{\sim}{\theta}}$ and $\bar{\underset{\sim}{\theta}}$ and expanding $\ln \bar{\theta}_i$ $(\ln(1-\bar{\theta}_i))$ about $\hat{\theta}_i$ $((1-\hat{\theta}_i))$ that our test statistic can be written as $T_1 = -2 \ln \Lambda_1 = \sum_{i=1}^{k-1} \left[\frac{\sum_{j=1}^{i} \hat{p}_j}{\alpha_i^2} + \frac{\hat{p}_{i+1}}{\nu_i^2} \right] \cdot n \cdot (\hat{\theta}_i - \bar{\theta}_i)^2$

where α_i is between $\bar{\theta}_i$ and $\hat{\theta}_i$ (and thus converges a.s. to θ_i) and ν_i is between $1-\bar{\theta}_i$ and $1-\hat{\theta}_i$ (and thus converges a.s. to $1-\theta_i$).

By employing arguments similar to those used in Theorem 2.2, we are led to the following theorem. (Note that we do not need to restrict $\underset{\sim}{p}_n$ to H_1.)

<u>Theorem 2.4</u>. If $\underset{\sim}{p}_n$ converges to $\underset{\sim}{p}$ and $\delta_{n,i}$ (as defined in (2.11)) converges to δ_i ($\pm\infty$ are possible values) for $i = 1,\cdots,k-1$, then T_1 is distributed asymptotically as

$$V = \sum_{i=1}^{k-1} [(Z_i + \delta_i) \wedge 0]^2$$

where $Z_1,\cdots,Z_{k-1}$ are independent $n(0,1)$ random variables.

We note that if $i^{-1}\sum_{1}^{i} p_j > (i+1)^{-1} \sum_{1}^{i+1} p_j$, then $\delta_{n,i} \to \infty$. In this case, the i^{th} term in (2.13) is zero and can be ignored, leading to the following theorem.

Theorem 2.5. If p satisfies H_1 then

$$\lim_{n\to\infty} P[T_1 \geq t] = \bar{\chi}_m^2(t)$$

where m is the number of subscripts, i, such that $i^{-1}\Sigma_{j=1}^{i} p_j = (i+1)^{-1}\Sigma_{j=1}^{i+1} p_j$; $i = 1,2,\cdots,k-1$. In addition,

$$\sup_{p\in H_1} \lim_{n\to\infty} P_p[T_1 \geq t] = \lim_{n\to\infty} P_{H_0}[T_1 \geq t]$$

$$= \bar{\chi}_{k-1}^2(t)$$

where $P_{H_0}[T_1 \geq t]$ is the probability of the event $[T_1 \geq t]$ computed under H_0.

We note that Theorems 2.2 and 2.4 imply that the likelihood ratio tests considered here are consistent in the sense that for $\underset{\sim}{p}$ lying in the region defined by the alternative hypothesis, the power function must converge to one.

3. Poisson Problem.

Suppose we have a random sample of size n from each of k Poisson populations having means $\lambda_1, \lambda_2, \cdots, \lambda_k$. Shaked (1979) found the maximum likelihood estimate of $\underset{\sim}{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_k)$ subject to the restriction H_1 requiring that $\underset{\sim}{\lambda}$ is lower starshaped:

$$H_1 : \lambda_1 \geq \frac{\lambda_1 + \lambda_2}{2} \geq \cdots \geq \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_k}{k} \geq 0. \tag{3.1}$$

This result can be found in a straightforward fashion using the results in Section 2 together with the fact that the conditional distribution of independent Poisson variables, given their sum, is multinomially distributed.

We first write the likelihood function in terms of the variables $\phi_1, \phi_2, \cdots, \phi_k$ where

$$(3.2) \qquad \phi_i = \lambda_i / \Sigma_{i=1}^{k} \lambda_i ;\ i = 1,2,\cdots,k-1 \quad \text{and} \quad \phi_k = \Sigma_{j=1}^{k} \lambda_i$$

$(\lambda_i = \phi_i \phi_k;\ i = 1,2,\cdots,k-1,\ \ \lambda_k = \phi_k - \Sigma_{i=1}^{k-1} \phi_i \phi_k)$. The restriction that $\underset{\sim}{\lambda}$ is starshaped is equivalent to requiring that $\underset{\sim}{\phi}$ is starshaped, or

$$(3.3) \qquad H_1' : \phi_1 \geq \frac{\phi_1 + \phi_2}{2} \geq \cdots \geq (k-1)^{-1} \Sigma_{i=1}^{k-1} \phi_i \geq k^{-1}$$

and these restrictions do not involve ϕ_k. The likelihood function is proportional to

$$(3.4) \quad \left[\left(\Pi_{i=1}^{k-1} \phi_i^{n\bar{x}_i} \right) \cdot \left(1 - \Sigma_{i=1}^{k-1} \phi_i \right)^{n\bar{x}_k} \right] \cdot \left[e^{-n\phi_k} \cdot \phi_k^{n\Sigma_{i=1}^{k} \bar{x}_i} \right]$$

where $\bar{x}_i$ is the mean of the sample from the i^{th} population; $i = 1,2,\cdots,k$. Because H_1 does not restrict ϕ_k, the two factors in brackets may be maximized independently. Using the results from Section 2 on the first factor and an easy analysis involving the derivative of the second factor, we obtain the restricted maximum likelihood estimates as follows:

$$(3.5) \qquad \bar{\phi}_i = \hat{\phi}_i \vee (1/i+1) : \quad i = 1,2,\cdots,k-1$$

where $\hat{\phi}_i = \bar{x}_i / \Sigma_{j=1}^{k} \bar{x}_j$, and $\bar{\phi}_k = \Sigma_{j=1}^{k} \bar{x}_j = \hat{\phi}_k$. (Note that $\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_k$ are the unrestricted maximum likelihood estimates of $\phi_1, \phi_2, \cdots, \phi_k$.) Using the invariance property of maximum likelihood estimation, we have the following theorem.

Theorem 3.1 (Shaked). The maximum likelihood estimates of $\lambda_1, \lambda_2, \cdots, \lambda_k$ subject to the restriction H_1 are given by

$$\bar{\lambda}_i = \left[\frac{\bar{x}_i}{\sum_{j=1}^k \bar{x}_j} \vee \frac{i}{i+1}\right] \cdot \sum_{j=1}^k \bar{x}_j \; ; \quad i = 1, 2, \cdots, k-1 \quad .$$

(3.6) and

$$\bar{\lambda}_k = \left(\sum_{j=1}^k \bar{x}_j\right)\left[1 - \sum_{i=1}^{k-1}\left(\frac{\bar{x}_i}{\sum_{j=1}^k \bar{x}_j} \vee \frac{i}{i+1}\right)\right] .$$

The likelihood ratio statistic for testing $H_0 : \lambda_1 = \lambda_2 = \cdots = \lambda_k$ against the alternative $H_1 - H_0$ can be written in terms of the ϕ's, as follows:

$$(3.7) \qquad \Lambda_{01} = \frac{(1/k)^{n\sum_{i=1}^k \bar{X}_i}}{\prod_{i=1}^{k-1} \bar{\phi}^{n\bar{X}_i} \left(1 - \sum_{i=1}^{k-1} \bar{\phi}_i\right)^{n\bar{X}_k}} .$$

If we let $S_{01} = -2 \ln \Lambda_{01}$ and let $Y = n\sum_{i=1}^k \bar{X}_i$, then given that $Y = y$, the joint conditional distribution of $y\hat{\phi}_1, y\hat{\phi}_2, \cdots, y(1-\sum_{j=1}^{k-1}\hat{\phi}_j)$ is multinomial with parameters y and $\lambda_1/\sum_{i=1}^k \lambda_i, \lambda_2/\sum_{i=1}^k \lambda_i, \cdots, \lambda_k/\sum_{i=1}^k \lambda_i$.

If we let $\underset{\sim}{\lambda}_n$ satisfying H_1 converge to $\underset{\sim}{\lambda}$ ($\lambda_i > 0$ for all i) such that

$$(3.8) \qquad \delta_{n,i} = \sqrt{n}\left(\sum_{j=1}^{i} \lambda_{n,j} \Big/ \sum_{j=1}^{i+1} \lambda_{n,j} - \frac{i}{i+1}\right)\left[\frac{(i+1)^3}{ik}\right]^{1/2} \to \delta_i \text{ (possibly } \infty)$$

and let $\bar{X}_{n,i}$ denote the corresponding independent sample means which occur in $S_{01}^{(n)} = -2 \ln \Lambda_{01}^{(n)}$, then using the Dominated Convergence Theorem:

$$\lim_{n\to\infty} P(S_{01}^{(n)} \geq t) = \lim_{n\to\infty} E(P(S_{01}^{(n)} \geq t \mid Y_n))$$

$$= E[\lim_{n\to\infty} P(S_{01}^{(n)} \geq t \mid Y_n)] = E[P(U \geq t)]$$

$$= P[U \geq t]$$

where U is distributed as in Theorem 2.2. Thus S_{01} is distributed asymptotically exactly as T_{01} in Section 2 (with $p_{n,i}$ replaced by $\lambda_{n,i}$ in the $\delta_{n,i}$'s). In particular, if H_0 is true, the asymptotic tail probabilities are given as in Theorem 2.3.

Theorem 3.2. If H_0 is true, then

$$\lim_{n\to\infty} P[S_{01} \geq t] = \bar{\chi}^2_{k-1}(t).$$

Let us now consider the problem of testing H_1 as a null hypothesis. The likelihood ratio Λ can be written

$$(3.9) \qquad \Lambda_1 = \frac{\left(\Pi_{i=1}^{k-1} \bar{\phi}_i^{n\bar{X}_i}\right)\left(1-\Sigma_{i=1}^{k-1} \bar{\phi}_i\right)^{n\bar{X}_k}}{\left(\Pi_{i=1}^{k-1} \hat{\phi}_i^{n\bar{X}_i}\right)\left(1-\Sigma_{i=1}^{k-1} \hat{\phi}_i\right)^{n\bar{X}_k}}$$

since $\hat{\phi}_k = \bar{\phi}_k$.

If we let $S_1 = -2 \ln \Lambda_1$, then the same type of reasoning used in the previous argument can be applied to show that

$$\begin{aligned}(3.10) \qquad \lim_{n\to\infty} P(S_1 \geq t) &= \lim_{n\to\infty} E(P(S_1 \geq t \mid Y)) \\ &= E[\lim_{n\to\infty} P(S_1 \geq t \mid Y)] \\ &= P[V \geq t]\end{aligned}$$

where V is defined in Theorem 2.4 and $\delta_{n,i}$ is defined as in (3.8). It follows from (3.10) that if $\underset{\sim}{\lambda}$ satisfies H_1, then the i^{th} term of V goes to zero if

$$i^{-1}\sum_{j=1}^{i}\lambda_j > (i+1)^{-1}\sum_{j=1}^{i+1}\lambda_j,$$

which leads to the following theorem.

<u>Theorem 3.3</u>. If $\underset{\sim}{\lambda}$ satisfies H_1, then

$$\lim_{n\to\infty} P_{\underset{\sim}{\lambda}}[S_1 \geq t] = \bar{\chi}^2_m(t)$$

where m is the number of distinct i such that

$$i^{-1}\sum_{j=1}^{i}\lambda_j = (i+1)^{-1}\sum_{j=1}^{i+1}\lambda_j.$$

Moreover

$$(3.11)\quad \sup_{\underset{\sim}{\lambda}\in H_1}\lim_{n\to\infty} P_{\underset{\sim}{\lambda}}[S_1 \geq t] = P_{H_0}[S_1 \geq t] = \bar{\chi}^2_{k-1}(t)$$

where $P_{H_0}(\cdot)$ is computed under the assumption that H_0 is true.

Note that Theorem 3.3 enables us to construct likelihood ratio tests of a particular size asymptotically when testing H_1 versus all other alternatives. Of course (3.8) and (3.10) assure us that our tests are asymptotically consistent in the sense that if $\underset{\sim}{\lambda}$ is in the region of the alternate hypothesis, the probability of rejecting the null hypothesis converges to one as $n \longrightarrow \infty$.

It should be noted that even though Shaked (1979) allows the more general starshaped ordering

$$\lambda_1 \geq \sum_1^2 w_j\lambda_j / \sum_1^2 w_j \geq \sum_1^3 w_j\lambda_j / \sum_1^3 w_j \geq \cdots \geq \sum_1^k w_j\lambda_j / \sum_1^k w_j \geq 0,$$

his restriction that the sample size from the i^{th} population be proportional to w_i effectively reduces the problem to the one considered earlier.

4. Concluding Remarks.

F. T. Wright called our attention to the work of Shaked (1979) after we had carried through much of the research in this paper. Actually our original analysis neglected the nonnegativity restriction and we termed the restriction "decreasing on the average." More specifically we should have termed it "decreasing on the average from the left" since $\theta_1 \geq \frac{\theta_1 + \theta_2}{2} \geq \frac{\theta_1 + \theta_2 + \theta_3}{3}$ is not equivalent to $\frac{\theta_1 + \theta_2 + \theta_3}{3} \geq \frac{\theta_1 + \theta_2}{2} \geq \theta_3$ (i.e., increasing on the average from the right).

It is clear that the restrictions increasing on the average from the right, increasing on the average from the left and decreasing on the average from the right can be handled by analysis similar to that in Sections 2 and 3.

The phrase "decreasing on the average" also calls to mind the restriction

$$H_2 : i^{-1} \sum_{j=1}^{i} \theta_j \geq (k-i)^{-1} \sum_{j=i+1}^{k} \theta_j \; ; \; i = 1,2,\cdots,k-1.$$

An equivalent way of stating H_2 is $i^{-1}\sum_{j=1}^{i}\theta_j \geq k^{-1}\sum_{j=1}^{k}\theta_j$; $i = 1,2,\cdots,k-1$. We note that the order restrictions specified in H_2 are less restrictive than those imposed by H_1 which in turn are less restrictive than $\theta_i \geq \theta_{i+1}$, $i = 1,\cdots,k-1$. In the multinomial setting, maximum likelihood estimates of p subject to H_2 and distribution theory for testing H_0 vs. H_2-H_0 and for testing H_2 as a null hypothesis can be found in Robertson and Wright (1980). Again, the asymptotic distribution is a chi-bar-square.

The 10%, 5% and 1% cutoff values of $\bar{\chi}^2(\cdot)$ for $k = 2,3,\cdots,15$ are given in Table 1.

Table 1.

Critical values for $\bar{\chi}_{k-1}(t) = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{1}{2}\right)^{k-1} P[\chi_{\ell}^{2} \geq t]$

k \ α	.10	.05	.01
3	2.95	4.23	7.28
4	4.01	5.44	8.77
5	4.95	6.50	10.02
6	5.84	7.48	11.18
7	6.67	8.41	12.26
8	7.48	9.29	13.31
9	8.26	10.15	14.29
10	9.02	10.99	15.29
11	9.76	11.79	16.21
12	10.49	12.59	17.12
13	11.22	13.38	18.01
14	11.93	14.15	18.91
15	12.63	14.91	19.78

5. Acknowledgement.

F.T. Wright suggested the investigation which led to this paper. Our original derivation of the results concerning Poisson parameters was more complicated than the analysis presented here. The suggestion to use the multinomial results was made by Moshe Shaked. The computer program which generated the table of $\bar{\chi}^2$ cutoff values was written by Daniel M. Flach.

REFERENCES

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). Statistical Inference under Order Restrictions. Wiley: New York.

Bartholomew, D.J. (1959). A test of homogeneity for ordered alternatives, Biometrika 46, 36-48.

Billingsley, Patrick (1968). Convergence of probability measures. Wiley: New York.

Chacko, V.J. (1966). Modified chi-square test for ordered alternatives, Sankhya Ser. B, 28, Parts 3 and 4, 185-190.

Kepner, James L. (1979). Tests using the null hypothesis of bivariate symmetry. Ph.D. thesis, University of Iowa.

Robertson, Tim (1978). Testing for and against an order restriction on multinomial parameters, Journal of the American Statistical Association, 73, 197-202.

Robertson, Tim and Wegman, Edward J. (1978). Likelihood ratio tests for order restrictions in exponential families, Ann. Statist. 6, 485-505.

Robertson, Tim and Wright, F.T. (1980). On measuring the conformity of a parameter set to a trend, with applications. University of Iowa, Department of Statistics Technical Report No. 71 (submitted).

Serfling, Robert J. (1980). Approximation Theorems of Mathematical Statistics. Wiley: New York.

Shaked, Moshe (1979). Estimation of starshaped sequences of Poisson and normal means, Ann. Statist. 7, 729-741.